DOCUMENT RESUME

ED 332 175 CS 212 121

AUTHOR Warren, John
TITLE Hypertext: A Tool for Academic Exploration.
PUB DATE Oct 89
NOTE 26p.; Paper presented at the Annual Meeting of the Humanities and Technology Association (13th, Atlanta, GA, October 19-20, 1989).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142) -- Historical Materials (060)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Structures; Computers; Context Clues; Databases; Futures (of Society); *Hypermedia; *Information Processing; *Language Processing; *Man Machine Systems; *Researchers; Research Methodology; Technological Advancement; Writing Processes
IDENTIFIERS Computer Mapping; Writing Contexts

ABSTRACT

This paper presents background information about the development and functions of hypertext and predicts future changes in converting text into hypertext. The paper describes hypertext as a computer system that allows the interconnection of information in a way that is "transparent" (does not intrude on or interfere with the reader's concentration) and is "intuitive" (functioning in the way humans are used to working instead of forcing them into a rigid framework). The paper compares the basic functional unit of the hypertext system to the multidimensional constructions of the human mind because both are fragmentary in effect and do not function in a linear mode. In speculating about future developments, the paper suggests, however, that hypertext should not replace conventional linear exposition, but that the skill of linking fragments will be of equal importance in writing and that the writer will be thought of as a "knowledge engineer." (Four figures are included.) (KEH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED332175

Hypertext:

A tool for academic exploration

John Warren

University of Southern Mississippi

BEST COPY AVAILABLE

CS212121

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Warren

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The author is torn between an artistic need to build logically, adding item upon item, until the reader is prepared for and, indeed, has almost intuitively developed the concept of hypertext and the logical recognition that many of his audience may be impatiently waiting for a definition of the term. In recognition that the academic paper is not a mystery novel, he will begin with a definition of hypertext. However, the reader is cautioned; the map is not the territory; violin music can be defined as "The drawing of the stretched hair of a horse across the dried entrails of a cat." Knowing the definition and appreciating the thing may be two different things.

To use the words of Ted Nelson (1967, p. 195), who created the term hypertext, hypertext or non-linear text differs "from ordinary text in its sequencing (it may branch into trees or networks), its organization (it may have multiple levels of summary and detail), its mode of presentation (it may contain moving or manipulable illustrations, moving or flashing typography), and so on. At last, with the appearance of computer display and

appropriate backup systems, such texts are possible and practical."

Insert Figure 1 about here

Carl Franklin (1989) calls it "an associative information management system" but admitted,

> "Defining hypertext is much like describing beauty: part of it is inherent in the eye of the beholder. With this in mind let's examine what we know about hypertext.
>
> "It is associative in that a single word or an entire document can be linked to another word or group of words in any way that is meaningful to the user.
>
> "Each link represents information (as imprecise as that may sound).
>
> "Browsers and other mechanisms manage the associative information represented by these links very differently than a database management system (DBMS) which relates information in a highly structured manner. (p. 37)"

Conklin (1987) approached a definition with trepidation.

> "It is tempting to describe the essence of hypertext as its ability to perform high-speed, branching transactions on textual chunks. But this is a little like describing the essence of a great meal by listing its ingredients. Perhaps a better description would focus on hypertext as a computer-based medium for thinking and communication.
>
> "The thinking process does not build on new ideas one at a time, starting with nothing and turning out each idea as a finished pearl. Thinking seems rather to proceed on several fronts at once, developing and rejecting ideas at different levels and on different points in parallel, each idea depending on and contributing to others. (p.32)"

The author would describe rather than define hypertext. It is a computer system that allows the interconnection of information in a way that is "transparent" (does not intrude on or interfere with the reader's concentration) and is "intuitive"

(functioning in the way humans are used to working instead of forcing them into a rigid framework).

Hypertext exists and is in limited use in a number of academic areas. In the interest of conciseness and because of the linear nature of the academic presentation, this paper will not deal in detail with specific, present-day applications but will, instead, present a background of hypertext and attempt to predict developments. As such, the latter part of the paper may be seen as highly speculative particularly when it deals with social changes brought about by the conversion of text into hypertext.

Again, in the interests of conserving time, this paper will not extend its investigations into hypermedia, a concept which extends the hypertext concept into linking text, sound, film, slides and other qualitatively different media. This is not an attempt to denigrate the importance of this concept. The author feels that hypermedia is the inevitable extension but recognizes that the concepts discussed in this paper will be as relevant to a hypermedia- as to a hypertext-document-production system.

DEVELOPMENT

Probably since the invention of writing there have been those who have voiced the grievance, "This just isn't the way my mind works." Scribe and clay, pen and paper, and even keyboard and CRT have seemed to impose a linear structure that was somehow alien to the way the human mind functions. Of course, writers have developed various strategies to provide a sort of multidimensionality to the linear text. All academics are familiar with footnotes, bibliographic references, and intersectional references (e.g. see page 26). Journalists use sidebars to go into detail about concepts whose exposition in the main story would obstruct the linear flow. Authors have long used the parenthetical comment and the flashback as non-linear tools for story telling. However, the immutability of print has mitigated against any expansion of these techniques.

Recent developments in cognitive psychology support the contention that the mind does not function in a linear mode. Knowledge structures are seen as multidimensional constructions consisting of ideas or schema that are linked by

the attributes related to each schema. Because the knowledge structure consists of schemas for all attributes the schemas are said to be embedded within each other.

Some cognitive psychologists have postulated what they call an active structural network, which is composed of schema and linked by ordered, labeled relationships. In this network learning took place when new material was brought in and integrated with existing schema. Obvious, this is quite different from the linear addition of material that takes place during writing or reading.

Of course, during the initial entry into a subject field, there are so few schema that the relationships develop easily despite the mode of introduction. This is particularly true when the domain is simple and well structured. However, as Spiro (1988, p. 10) put it, "In an ill-structured domain, emphasis must be shifted from retrieval of intact, rigid, precompiled knowledge structures, to assembly of knowledge from different conceptional and precedent case sources to adaptively fit the

situation at hand. In this arena, linear text has little to recommend it.

The development of the computer created the potential for a new form of text. As early as 1945, Vannevar Bush, then director of the U.S. Office of Scientific Research and Development conceived of what he called a "Memex" and wrote:

"Consider a future device in which an individual stores all his books, records, and communications, and which may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory." He postulated a machine that would use association for data recovery rather than the alphabetic or numerical sorts that Hollerith cards had already made popular. Bush wrote, "With one item in its grasp, it snaps instantly to the next item that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain."

Because the digital computer was in its infancy, Bush postulated an analog device using light sensitive cells and microfilm readers. However, he accurately presented two concepts that

are central to hypertext: associative rather than hierarchial linking and the ability of the reader to interact with the media and add personal information to the file.

By 1967, the futurist Theodore Nelson could confidently predict combining a natural-language text with the computer's capability to provide an interactive branching and dynamic display to produce a non-linear text which could not be sensibly printed on a conventional page. He attached a name to this non-linear text: Hypertext and compared the creators of the then-current database managers to the early movie producers.

Nelson contended both echoed an earlier art. The early movie producer attempted to reproduce the experience of the playgoer while the data-base programmer essentially duplicated the familiar card file. Nelson scored these misunderstandings of the media and failures to recognize how their character differed from that of the medium that preceded them. Just as the camera had capabilities in advance of stage (e.g. special effects, display arrangements, and the ability to select and interconnect parts, computers had capabilities that

were in advance of static systems. As Nelson summed up his article,

> "The problem of Getting It Out of Our System, then, is not the problem of fishhook design for a document pool, or of creating a conversational black box with a narrow vocabulary. We must get out of our system the fixities of thinking and procedure that hold us back. (p. 209)"

FUNCTIONING

The basic unit of the hypertext system or document is the fragment. Raymond and Tompa (1988) define fragments as

> "pieces of content or structure which are both discrete and independent. By discrete we mean that the distinction between the components is explicit and well-defined; by independent we mean that the components are capable of standing in multiple relations with each other. Ideally, a hypertext node is a text fragment which contains a single, independent concepts" (p. 873).

A good example is a conventional footnote in a research paper.

The key to a hypertext system is its ability to allow the author or user to connect these fragments on an associative rather than a relational basis as is done in DBMS's (database management systems). This does not mean that a DBMS cannot simulate the workings or a hypertext system but the simulation will require considerable programming and will lack the elegant ease-of-use characteristic of hypertext.

Insert Figure 2 about here

In practice, this is done via "windows" and "hot buttons." Windows are blocks of text on the screen which can be moved and examined independently of other blocks of text on the same screen. This concept is familiar to anyone who has worked with the Macintosh computer, but may be less clear to those whose computing experiences have been on more conventional computers where information scrolls top to bottom and old text is banished to a never-never land as it disappears

from the top line. Hot buttons are words and groups of words, representing concepts, are set off in the hypertext document. In some computers, reverse display is used to accomplish the setting off; in others, bold face type, underlining, or other typographical signal is used. If while reading it the user finds one of these words or groups of particular interest, he or she can move the cursor to it and use a function key to signal the computer to activate the hypertext system.

The system should ideally work instantly; however, at the worst, in a few seconds, the linked document appears in a window. In some systems, the original block is reduced in size and moved to one corner of the screen; in others, it is overlaid but not removed by the new window. If there are several links available to the selected text, a menu window appears.

The new document will also contain set off words and phrases indicating further links that the user can follow. Of course, at any point, the user can retrace the steps back to the original document.

Using these tools, users can carry on what van der Merwe (1989, p. 87) calls "information reconnaissance." They will have huge reserves of information without having it thrust upon them. However, at any point, they can move from a condensed abstract to more detailed information whenever, in the pursuit of an interesting concept, they feel the need. The author, on the other hand, can present the main point of the document without having that point obscured by a torrent of supporting documentation.

Interestingly, the footnote was developed with just such an end in mind. However, the computer's ability to exhaustively search a database has resulted in such an increase in footnotes that they, now, are coming to pose the same kind of obstacle they were designed to avoid. In a recent issue of Journalism Quarterly, the author noticed a article (Johnson & Christ, 1988) whose first page consisted of one inch of text and 11.5 inches of footnotes!

Let's take part of a hypothetical hypertext session. The user might begin in a constitutional law text. As he reads a section on First Amendment

law, he notices a mention of the case Schenck v. United States. He moves the cursor over Schenck and activates the hypertext program. A window appears over the constitutional law text with the exact wording of the Court's decision in the case. While reading Justice Holmes' decision, he encounters a mention of the document which got Mr. Schenck arrested in the first place. With a keystroke, he can jump from the decision to the exact wording of that document.

Insert Figure 3 about here

Of course, our hypothetical searcher may not chose this particular route through the hypertext. He might have jumped from the text to the decision to biographical notes on Schenck to a text on 19th century socialism. The point is that a hypertext is not a passive activity.

Insert Figure 4 about here

Both the author (or knowledge engineer) and the reader share in creating a unique experience.

An added advantage is that the computer can save the "map" so that searcher can easily retrace his or her steps through the nodes or can direct others to do so.

FUTURE

Obviously, hypertext is not a replacement for linear exposition. Hypertext should not, or even could, replace the conventional linear document. The linear format is sanctioned by more than simple convention. Properly done, the flow of narrative is an important component of the art of writing. It is not difficult to imagine the destructive effects on literature if the reader is able to wander away from the main plot. Hamlet would be infinitely worse off if the reader, instead of following him to the battlements and the meeting with his father, was instead wandering through the construction techniques used in Danish castles. As Raymond and Tompa (1988, p. 873) put it, "We suggest the linear structure is a virtue achieved at some cost by the document's creator. Furthermore, a subtle treatment of a theme may communicate an idea more artistically (and hence more successfully) then any explicit purpose."

This is not to indicate that fiction cannot be produced in hypertext format. A recent science fiction novel, David's Sling by Marc Stiegler was simultaneously published in conventional paperback and in hypertext, but it is difficult to visualize the majority of fiction readers abandoning their easy chairs to hunch over their computers

However, research and scientific documents are admirably suited to hypertext. Not only is their purpose the communication of information but they are rarely written in a style that could be said to "artistically communicate an idea."

Even here, there will be resistance. Writers, by and large, have considerable pride in their wordsmithing. Hypertext calls for a quantum jump. The writing will still be important, but skill in linking the fragments (many of whom will not have been produced by the author) will be of equal importance. While writers are accustomed to giving credit through footnotes and attributed quotes, it can be seen that considerable pride is involved in the selection of the precisely applicable quote and an certain proprietary air is generated by that action. Will authors be able to claim the same

pride of authorship when much of the work consists of linking applicable documents? In short, will people be willing to attach the same pride to the title of "knowledge engineer" as to "author?"

Or, will knowledge engineer become a new intellectual profession -- perhaps growing from the ranks of librarians whose responsibilities have expanded exponentially in recent years. Clearly, the creation of hypertext documents will be a monumental task. This will be particularly true initially where all the links must be developed by the pioneering author. Later, authors can tap into the existing networks and use these links as part of their output.

One hopeful sign is the increasing number of documents available in an computer-readable format. These include such disparate titles as Bowker's Books in Print, Microsoft's Bookshelf (a compendium of ten works including Roget's Thesaurus, Bartlett's Familiar Quotations, the World Almanac, and the U.S. Zip Code Directory), the McGraw-Hill Concise Encyclopedia of Science and Technology as well as the Tibetan Book of the Dead and Steppenwolf.

Copyright and the appropriate payments due to an author will also present problems. The law has been slow to adapt to electronic publishing and hypertext may increase the existing problems by an order of magnitude. Some systems, for example Ted Nelson's Xanadu, carefully maintain copyright protection and incorporate ways to account for and pay royalties.

However, it is not completely clear what will be a copyrightable document in a hypertext world. One wit commented that to steal from one author was plagiarism; to steal from many was scholarship, but in truth, many scholarly pieces are primarily the drawing together of other peoples' writing. The spark of genius there is the being able to synthesize and combine -- making a whole that is greater than the sum of its parts. Today, this is clearly a work of writing, but what happens when the drawing together consists primarily of computer-generated links? Can that still stand as a work of creativity? Related to this is the value the academic power-structure places on the "publication" of a hypertext document. The development of the logical extension of the

academic journal, the electronic journal, has largely been stymied by just this question. Because many promotion and tenure committees question the legitimacy of an electronic journal, academics are reluctant to write for them. One academic phrased the situation thus: "Major authors won't write for the electronic journals because they are seen as low status; electronic journals are seen as low status because major authors don't write for them." Hypertext may find itself in just such a impasse.

Such questions need to be answered because the acceptance of hypertext is a social more than a technological problem.

Figure 1

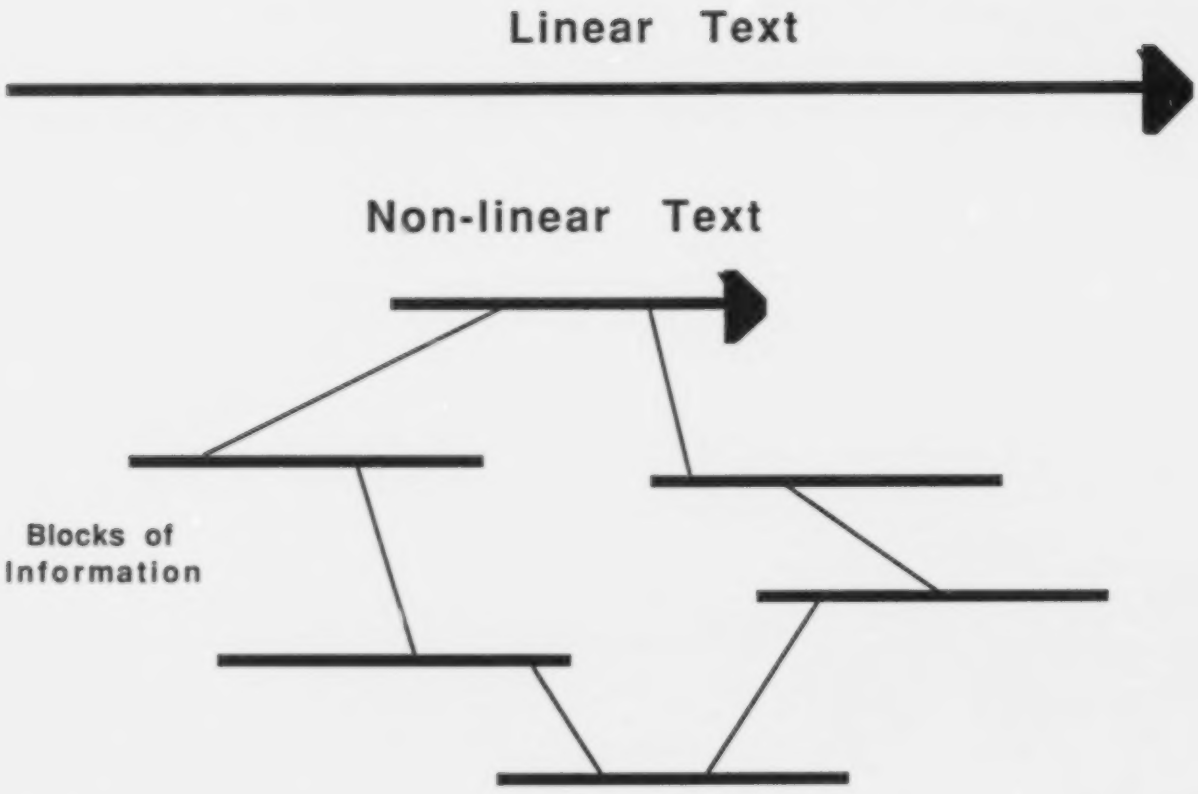


Figure 2

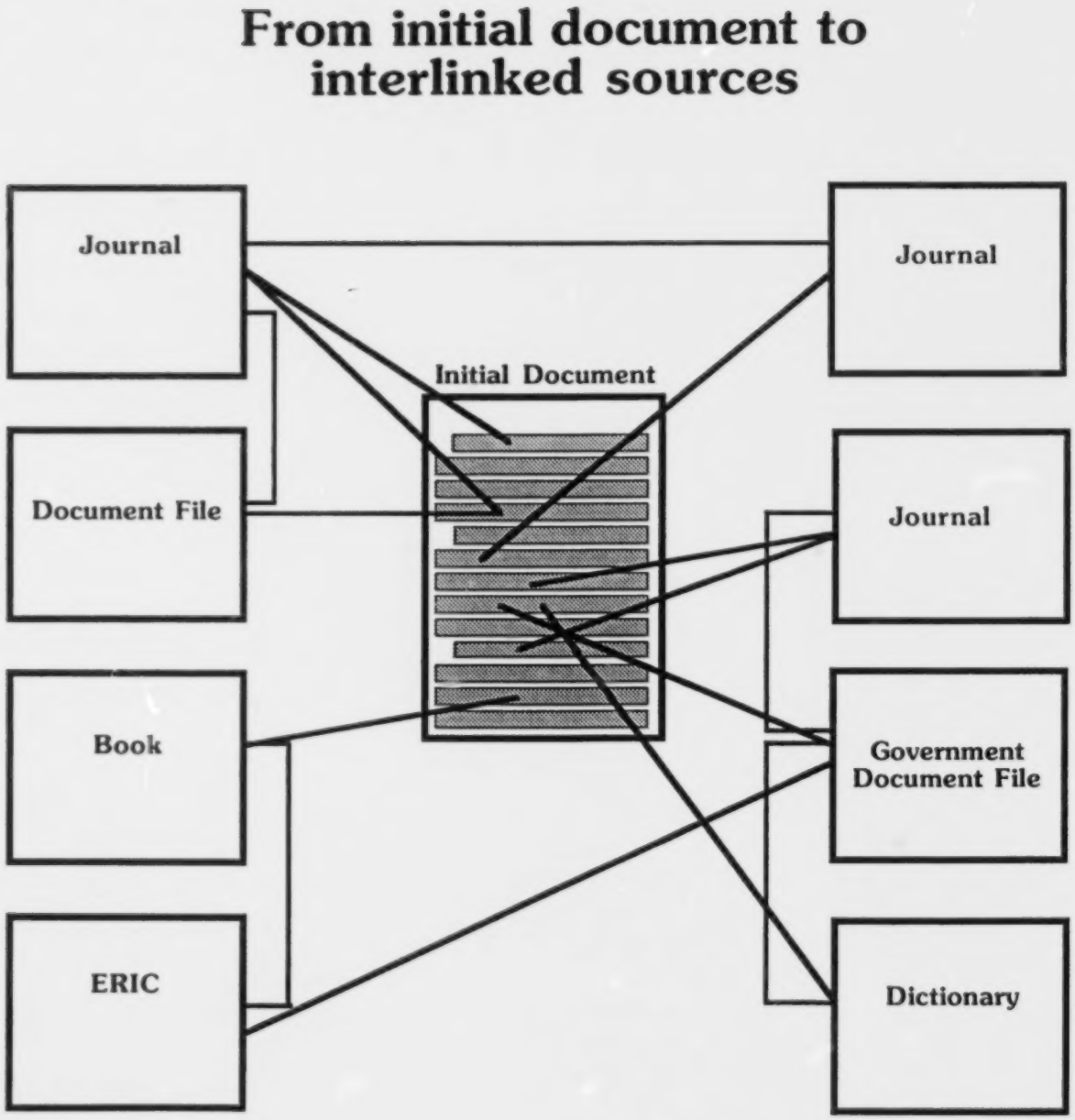

From initial document to interlinked sources
Journal
Document File
Book
ERIC
Initial Document
Journal
Journal
Government Document File
Dictionary

Figure 3

Sample of hypothetical Hypertext session

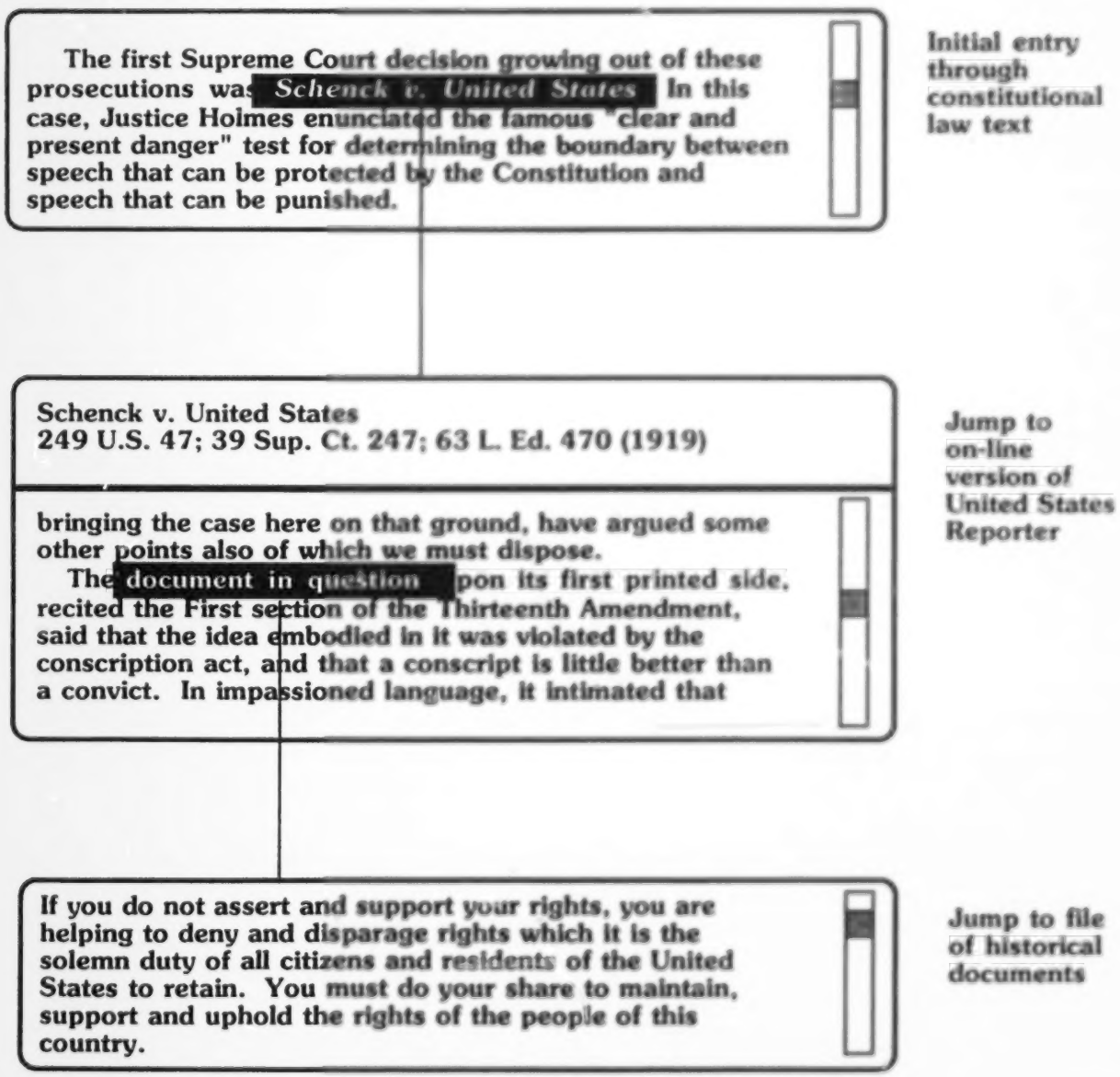
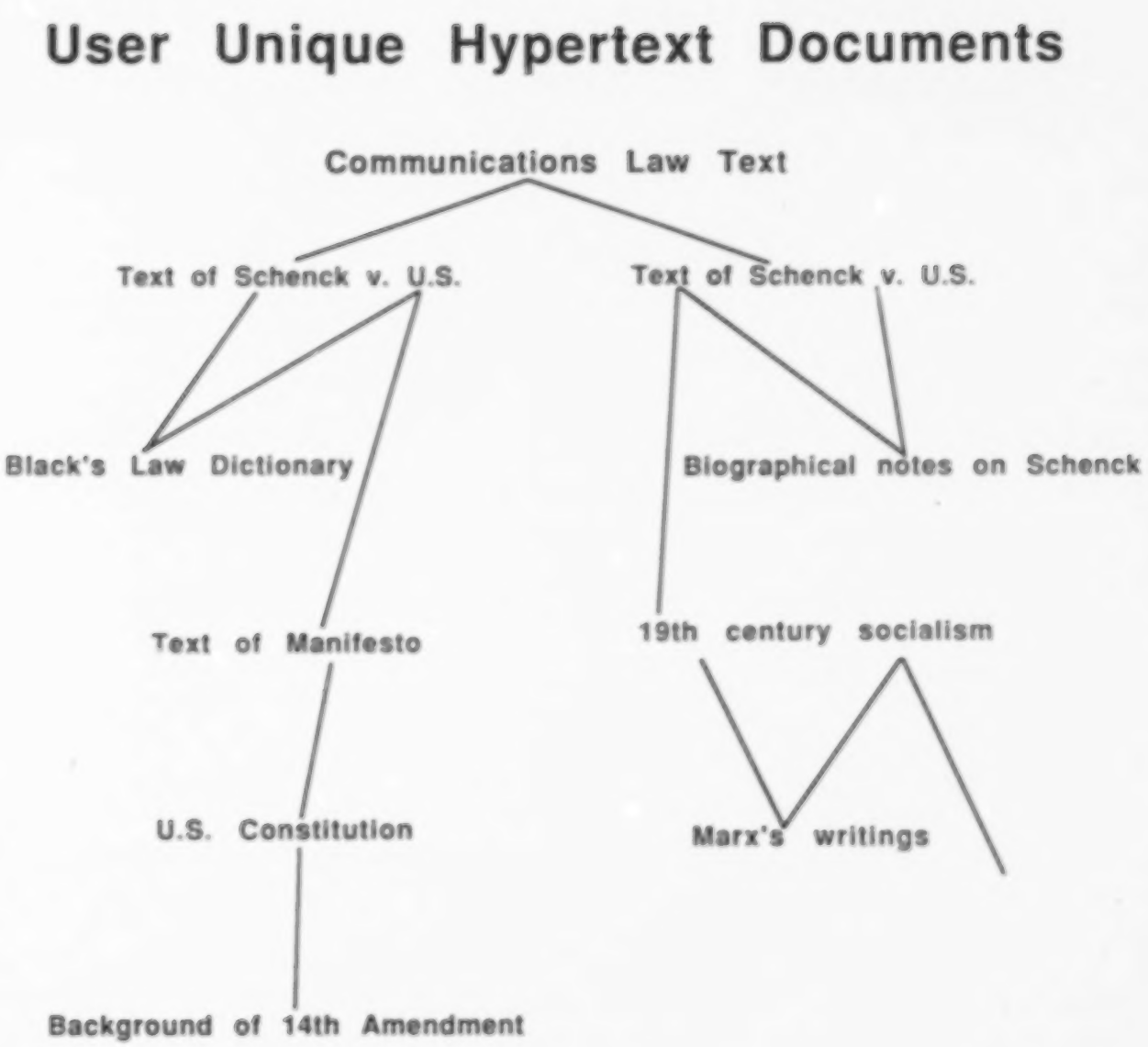

The first Supreme Court decision growing out of these prosecutions was *Schenck v. United States*. In this case, Justice Holmes enunciated the famous "clear and present danger" test for determining the boundary between speech that can be protected by the Constitution and speech that can be punished.
Initial entry through constitutional law text
Schenck v. United States
249 U.S. 47; 39 Sup. Ct. 247; 63 L. Ed. 470 (1919)
Jump to on-line version of United States Reporter
bringing the case here on that ground, have argued some other points also of which we must dispose.
The document in question upon its first printed side, recited the First section of the Thirteenth Amendment, said that the idea embodied in it was violated by the conscription act, and that a conscript is little better than a convict. In impassioned language, it intimated that
If you do not assert and support your rights, you are helping to deny and disparage rights which it is the solemn duty of all citizens and residents of the United States to retain. You must do your share to maintain, support and uphold the rights of the people of this country.
Jump to file of historical documents

Figure 4


User Unique Hypertext Documents
Communications Law Text
Text of Schenck v. U.S.
Text of Schenck v. U.S.
Black's Law Dictionary
Biographical notes on Schenck
Text of Manifesto
19th century socialism
U.S. Constitution
Marx's writings
Background of 14th Amendment

References

Bush, V. (1945, July). As we may think. Atlantic Monthly, pp. 101-108

Claassen, W.T., & Bothma, T.J.D. (1988). Structuring diverse types of information in hypertext: The case of Biblical information. South Africa Online Information 88. 12th International Online Information Meeting (vol.1 pp.83-90). London, UK

Conklin, J. (1987). Hypertext: An introduction and survey. Computer, 20(9), 17-40.

Dennis, S.F. (1967). The design and testing of a fully automatic indexing-searching system for documents consisting of expository text. In G. Schecter (ed.). Information retrieval: A critical view (pp. 67-94). Washington, DC: Thompson Book Company.

Johnson, S., & Christ, W. G. (1988). Women through Time: Who gets covered? Journalism Quarterly, 65(4), 889-897.

Nelson, T.H. (1967). Getting it out of our system. In G. Schecter (ed.). Information retrieval: A critical view (pp. 191-210). Washington, DC: Thompson Book Company.

Raymond, D.R., & Tompa, F.W. (1988). Hypertext and the Oxford English Dictionary. Communications of the ACM, 31, 871-879.

Spiro, R.J. (1988). Cognitive flexibility theory: Advanced knowledge Acquisition in Ill-structured domains. (Report No. 441). *** (ERIC Document Reproduction Service No. ED302821).

Stover, D. (1989, May). Hypermedia. Popular Science, p. 122-124, 160.

van der Merwe, D.P. (1988). Annotating literary texts with hypertext. South Africa Online Information 88. 12th International Online Information Meeting (vol. 1 pp. 239-47). London, UK